

บทปริทัศน์บทความเรื่อง Choice–Agreement Index and Its Application to Item Analysis

ศิริชัย กาญจนวาสี

บทความเรื่อง “Choice–Agreement Index and Its Application to Item Analysis” โดยศาสตราจารย์ ดร.ธีระ อาชวเมธี ได้เสนอแนวคิดและการพัฒนาดัชนีสำหรับวัดความสอดคล้องของการเลือกตัวเลือก (Choice–Agreement Index) เพื่อใช้กับข้อมูลที่ตัวเลือกวัดอยู่ในมาตรานามบัญญัติ (Nominal Scale) ดัชนีความสอดคล้องของการเลือกมีสูตรที่พัฒนาบนพื้นฐานของจำนวนคู่ที่สอดคล้องกันของการเลือกและจำนวนคู่ที่ไม่สอดคล้องกัน การคำนวณค่าดัชนีดังกล่าวมีพิสัยจาก 0 ถึง 1 ดัชนีความสอดคล้องนี้อาจนำไปใช้วิเคราะห์การแจกแจงตัวเลือก ตลอดจนประยุกต์ใช้ในการวิเคราะห์ข้อสอบแบบหลายตัวเลือก (Multiple Choice Test) สำหรับศึกษาลักษณะการเลือกตัวลวงของผู้ตอบ เพื่อสะท้อนคุณภาพของกลุ่มตัวลวงได้เช่นกัน

จากแนวคิดและการพัฒนาดัชนีดังกล่าว ผู้ปริทัศน์มีความคิดเห็น ข้อวิจารณ์ ข้อสังเกต และ ข้อเสนอแนะ ดังต่อไปนี้

1. การนำเสนอแนวคิดของการพัฒนาดัชนีความสอดคล้องของการเลือก

ผู้เขียนบทความได้เสนอแนวคิดของการพัฒนาดัชนีสำหรับวัดความสอดคล้องของการเลือกตัวเลือก I_{CA} ซึ่งอยู่ในมาตรานามบัญญัติได้อย่างมีลำดับขั้นตอนพร้อมทั้งยกตัวอย่างประกอบได้อย่างชัดเจน ตั้งแต่สูตรที่ 1 จนถึงสูตรที่ 12 ซึ่งเป็นสูตรสุดท้าย I_{CA} อยู่บนพื้นฐานของสัดส่วนความสอดคล้องระหว่างจำนวนคู่ที่ตอบได้สอดคล้องกัน (agreement pairs) จากจำนวนคู่ที่เป็นไปได้ทั้งหมด (all possible pairs) ซึ่งสามารถทำความเข้าใจได้ไม่สลับซับซ้อน

2. การคำนวณดัชนีความสอดคล้องของการเลือก และการทดสอบนัยสำคัญทางสถิติ

2.1 สูตรการคำนวณดัชนีความสอดคล้องของการเลือกที่ผู้เขียนบทความได้พัฒนาขึ้นมีสูตรที่ (10) น่าจะเป็นสูตรที่ทำความเข้าใจและคำนวณได้ง่ายกว่าสูตรอื่น ดังตัวอย่าง

$$I_{CA} = \sqrt{1 - \frac{2K \sum_{i < j} n_i n_j}{(K-1)N^2}}$$

- เมื่อ N = จำนวนผู้ตอบ
 K = จำนวนตัวเลือก
 i = สัญลักษณ์แทนตัวเลือก จาก 1 ถึง $K-1$
 j = สัญลักษณ์แทนตัวเลือก จาก 2 ถึง K

สมมติในการตอบข้อสอบข้อหนึ่งที่มี 5 ตัวเลือก จำนวนผู้ตอบ 100 คน เป็นดังตาราง

ตาราง แสดงการแจกแจงความถี่ในการเลือกตัวเลือกของข้อสอบ

ตัวเลือก	ความถี่	สัดส่วน
1 (ก)	25	0.25
2 (ข)	20	0.20
3 (ค)	10	0.10
4 (ง)	5	0.05
5 * (จ)	40	0.40
รวม	100	1.00

หมายเหตุ * คำตอบถูก

การคำนวณดัชนีความสอดคล้องของการเลือกตัวลวง

$$I_{CA(\text{ตัวลวง})} = \sqrt{1 - \frac{2(4)[(25 \times 20) + (25 \times 10) + (25 \times 5) + (20 \times 10) + (20 \times 5) + (10 \times 5)]}{(4-1)(60^2)}}$$

$$= 0.30$$

ดังนั้นดัชนีความสอดคล้องของการเลือกตัวลวงสำหรับข้อสอบนี้ หรือ “ดัชนีความลวงร่วม” มีค่าเท่ากับ 0.30

2.2 การทดสอบนัยสำคัญทางสถิติของความแตกต่างระหว่างดัชนีความสอดคล้องของกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระจากกัน ผู้เขียนบทความเสนอว่าถ้า K ซึ่งทำเป็นจำนวนตัวเลือกมีค่ามาก (หรือขนาดใหญ่) ให้ใช้การทดสอบเอฟ (F-test) ที่องศาของความเป็นอิสระ $K-1$ และ $K-1$ แต่ถ้า K มีค่าน้อย (หรือขนาดเล็ก) ให้ใช้การทดสอบไคสแควร์ (χ^2 - test) (ที่องศาความเป็นอิสระ $K-1$, ผู้ปริทัศน์) โดยมีกลุ่มผู้ตอบจำนวนมากเพื่อให้ความถี่ที่คาดหวังของแต่ละเซลล์มีค่าไม่ต่ำกว่า 10 (กรณี 2 ตัวเลือก) หรือมีค่าไม่ต่ำกว่า 5 (กรณีมากกว่า 2 ตัวเลือก)

ในทางปฏิบัติคงจะมีปัญหาในการตัดสินใจเลือกใช้สถิติทดสอบว่าขนาด K เท่าใดที่ถือว่าใหญ่พอจึงต้องการการการศึกษาวิจัยเพิ่มเติม นอกจากนี้กรณีที่เลือกใช้การทดสอบไคสแควร์ ถ้าปรากฏมีบางตัวเลือกที่ไม่มีใครเลือกเลยจะทำให้เกิดปัญหาเซลล์ว่างเปล่า (ความถี่เป็น 0) ซึ่งสามารถส่งผลกระทบต่อความน่าเชื่อถือของผลการทดสอบได้

3. การประยุกต์ใช้สำหรับการวิเคราะห์ข้อสอบ (Item Analysis)

การวิเคราะห์ข้อสอบโดยทั่วไปตามทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory) มีการวิเคราะห์คุณภาพของข้อสอบทั้งตัวเลือกที่เป็นคำตอบถูกและตัวลวง โดยคำตอบถูกมีการคำนวณดัชนีความยากของข้อสอบ (p) และดัชนีอำนาจจำแนกของข้อสอบ (r) ส่วนตัวลวงมีการคำนวณดัชนีเป็นรายตัวลวง ได้แก่ดัชนีการเลือกตัวลวง (p_w) และดัชนีอำนาจจำแนกของตัวลวง (r_w) โดยดัชนีทั้งสองของตัวลวงควรมีค่าไม่ต่ำกว่า 0.05 (ศิริชัย กาญจนวาสี, 2544)

สิ่งที่ขาดหายไปในการวิเคราะห์ข้อสอบตามทฤษฎีการทดสอบแบบดั้งเดิมคือ “ดัชนีความลวงร่วม” ซึ่งเป็นความสามารถในการลวงของกลุ่มตัวลวง (ผู้เขียนบทความนี้ใช้คำว่า *collective distractibility index*) ดัชนีนี้จะเป็นประโยชน์สำหรับการพิจารณาคุณภาพของกลุ่มตัวลวง ถ้าดัชนีความลวงร่วมของข้อสอบข้อใดก็ตามมีค่าต่ำ แสดงว่ากลุ่มตัวลวงของข้อสอบนั้นมีเส้นแห่งความเป็นตัวลวงในระดับสูง ในทางตรงกันข้าม ถ้าดัชนีความลวงร่วมของข้อสอบมีค่าสูงแสดงว่าเส้นแห่งของกลุ่มตัวลวงอยู่ในระดับต่ำ นอกจากนี้ควรได้มีการศึกษาวิจัยต่อเกี่ยวกับ *ดัชนีอำนาจจำแนกร่วม* ของกลุ่มตัวลวง

4. การศึกษาความสัมพันธ์ระหว่างดัชนีความยากกับดัชนีความลวงร่วม

ผู้เขียนบทความนี้ได้ศึกษานำร่องถึงความสัมพันธ์ระหว่างดัชนีความยากกับดัชนีความลวงร่วม โดยใช้ข้อมูลจากนักศึกษา 25 คน ข้อสอบ 27 ข้อ พบว่าสัมประสิทธิ์สหสัมพันธ์มีค่าเท่ากับ +0.49 ($p < 0.01$) ค่าที่ได้ยังคงห่างไกลจากการเป็นผลสรุป เนื่องจากเมื่อพิจารณาจากผลการวิเคราะห์รายข้อจะเห็นได้ว่าข้อสอบส่วนใหญ่ค่อนข้างง่าย ความยากง่ายของข้อสอบ (ตัวถูก) และความสามารถในการสร้างตัวลวงของผู้ออกข้อสอบ ย่อมจะมีผลต่อวินิจฉัยในการเลือกตัวเลือกของผู้ตอบซึ่งส่งผลต่อดัชนีความลวงร่วม ถ้าข้อสอบค่อนข้างง่าย (ดัชนีความยากมีค่าสูง) ผู้เลือกตัวลวงจะมีจำนวนน้อยและเป็นกลุ่มที่มีความสามารถต่ำ การเลือกตัวลวงจะสอดคล้องกันเพียงใดคงขึ้นกับคุณภาพของการสร้างตัวลวงแต่ละตัวด้วย แต่ถ้าข้อสอบค่อนข้างยาก (ดัชนีความยากมีค่าต่ำ) ผู้เลือกตัวลวงจะมีจำนวนมากขึ้น และเป็นกลุ่มที่มีความสามารถคละกัน การเลือกกลุ่มตัวลวงจะสอดคล้องกันเพียงใดก็คงขึ้นกับคุณภาพของการสร้างตัวลวงแต่ละตัวด้วยเช่นกัน ประกอบกับการคำนวณค่าดัชนีคุณภาพข้อสอบตามแนวทฤษฎีการทดสอบแบบดั้งเดิมมีการผันแปรไปตาม

ลักษณะของกลุ่มผู้สอบ จึงเป็นการยากที่จะได้ข้อสรุปที่เป็นนัยทั่วไปจากการศึกษาดังกล่าวข้างต้น นอกจากนี้จะสร้างสถานการณ์ (simulation) ให้ครอบคลุมเงื่อนไขต่าง ๆ อย่างรัดกุม

จากปัจจัยต่าง ๆ ที่กล่าวมาดั่งนั้นสัมประสิทธิ์สหสัมพันธ์ระหว่าง ดัชนีความยากของข้อสอบกับดัชนีความลงร่วมของกลุ่มตัวลองไม่น่าจะมีค่าสูง และมีความเป็นไปได้ที่จะมีค่าได้ทั้งทิศทางบวกและลบ

5. การพัฒนาโปรแกรมคอมพิวเตอร์

เมื่อได้มีการศึกษาวิจัยพัฒนาสูตรการคำนวณอย่างครอบคลุมการบ่งชี้คุณภาพตัวลอง และดัชนีอื่นที่เกี่ยวข้องจนเป็นที่ยอมรับในระดับหนึ่งแล้ว ควรมีการพัฒนาโปรแกรมคอมพิวเตอร์เพื่ออำนวยความสะดวกสำหรับการใช้ประโยชน์ให้กว้างขวางต่อไป

กล่าวโดยสรุป ผู้เขียนบทความได้เสนอแนวคิดและการพัฒนาสูตรการคำนวณดัชนีความสอดคล้องของการเลือกตัวเลือกที่มีลักษณะเป็นมาตรฐานบัญญัติได้อย่างมีลำดับขั้นตอนและน่าสนใจ ดัชนีความสอดคล้องของการเลือกอยู่บนพื้นฐานของการคำนวณจำนวนคู่ที่ผู้ตอบตอบได้สอดคล้องกัน จากจำนวนคู่ที่เป็นไปได้ทั้งหมด นอกจากนี้ผู้เขียนบทความยังได้เสนอวิธีทดสอบนัยสำคัญของความแตกต่างของดัชนีความสอดคล้องระหว่างกลุ่มผู้ตอบ ในการประยุกต์ใช้กับการวิเคราะห์ข้อสอบตามแนวทฤษฎีการทดสอบแบบดั้งเดิม ทำให้สามารถคำนวณดัชนีความลงร่วมของกลุ่มตัวลองซึ่งเปิดประตูสู่การวิจัยเพื่อสร้างเสริมองค์ความรู้เกี่ยวกับการวิเคราะห์ข้อสอบตามแนวทฤษฎีดั้งเดิมได้อีกหลาย ๆ ประเด็น

เอกสารอ้างอิง

ศิริชัย กาญจนวาสี. (2544). **ทฤษฎีการทดสอบแบบดั้งเดิม** (Classical Test Theory). (พิมพ์ครั้งที่ 4). กทม: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Archwamety, Teara. (2003). **Choice–Agreement Index and Its Application to Item Analysis**. University of Nebraska at Kearney.